Iratxe Berriozabal Campillo.
21.XI.96.

# CAPÍTULO IV

# PRUEBAS DE HIPÓTESIS

4.1. Introducción

4.2. Pruebas de hipótesis sobre la media poblacional

4.3. Pruebas de hipótesis sobre la varianza poblacional

4.4. Pruebas de hipótesis que involucran a dos poblaciones

4.5. Potencia de una prueba de hipótesis

4.6. Cálculo del tamaño de muestra

**TEXTO:**

SPIEGEL, M. R.: *Estadística*, Madrid, McGraw-Hill, 1990; capítulo 10, págs. 223-250.

**BIBLIOGRAFÍA COMPLEMENTARIA:**

BERENSON, M. L. y LEVINE, D. L.: *Estadística para administración y economía*, Madrid, Interamericana, 1987; capítulos 9 y 10.

CANAVOS, G. C.: *Probabilidad y estadística*, Madrid, McGraw-Hill, 1988; capítulo 9.

MILLS, R. L.: *Estadística para economía y administración*, Madrid, McGraw-Hill, 1980; capítulo 7.

PRUEBAS DE HIPÓTESIS: EJEMPLOS

**EJEMPLO 1:**

Un auditor desea probar el supuesto de que el valor promedio de todas las cuentas por cobrar en una empresa determinada es de 260.000 ptas., para lo cual extrae una muestra aleatoria de n = 36, y calcula su media muestral.

¿Cuáles serían las hipótesis nula y alternativa?

**EJEMPLO 2:**

En el **ejemplo 1** determinar los valores críticos de la media muestral, para probar la hipótesis con un nivel de significación del 5%, sabiendo que la desviación estándar de las cuentas por cobrar es: $\sigma$ = 43.000 ptas.

**EJEMPLO 3:**

En los **ejemplos anteriores**, supongamos que la media muestral es $\overline{x}$ = 240.000 ptas. ¿Cómo determinar si se debe aceptar o rechazar la hipótesis nula?

**EJEMPLO 4:**

Supongamos que el auditor de los **ejemplos anteriores** parte de la hipótesis nula de que la media de todas las cuentas por cobrar es de 260.000 ptas. al mes. Si la media muestral es de 240.000 ptas., probar dicha hipótesis a un nivel de significación del 5%.

**EJEMPLO 5:**

En el **ejemplo 4**, teniendo en cuenta que $\sigma$ = 43.000 ptas. y n = 36, calcular la probabilidad de error de tipo II.

**EJEMPLO 6:**

Con referencia al **ejemplo 5**, determinar la potencia de la prueba.

**EJEMPLO 7:**

Si el auditor de los **ejemplos anteriores** considera que la diferencia entre su hipótesis nula (que el promedio de cuentas por cobrar sea al mes de al menos 260.000 ptas.), y el valor obtenido en la muestra $x$ = 240.000 ptas. es considerable, y considerando que el error de tipo I ($\alpha$) es de 0,05 y el de tipo II ($\beta$) es de 0,10: ¿cuál sería el tamaño de la muestra que debe extraerse como mínimo para llevar a cabo esta prueba?

**EJEMPLO 8:**

Se plantea la hipótesis nula de que la vida útil promedio de los focos de una marca A, es de al menos 4.200 horas. La vida promedio para una muestra aleatoria de n = 10 focos es de $\overline{x}$ = 4.000 horas, con desviación estándar muestral de S = 200 horas. Se supone que la vida útil de los focos tiene una distribución normal. Probar la hipótesis nula al nivel de significación del 5%.

**EJEMPLO 9:**

El salario promedio mensual para una muestra aleatoria de tamaño $n_A$ = 30 empleados de una empresa A, resulta ser $\overline{X}_A$ = 180.000 ptas., con $S_A$ = 14.000 ptas. En otra empresa B, una muestra aleatoria de $n_B$ = 40 empleados, tiene un salario promedio $\overline{X}_B$ = 170.000 ptas. con $S_B$ = 10.000 ptas. se supone que las desviaciones estándar de las dos poblaciones de salarios sean iguales. Probar la hipótesis de que no existe diferencia entre los salarios promedio mensuales de las dos empresas, a un nivel de significación del 5%.

**EJEMPLO 10:**

Un fabricante de automóviles obtiene datos del rendimiento de gasolina para una muestra aleatoria de n = 10 automóviles, utilizando gasolina con y sin un determinado aditivo. Si se utilizan los mismos conductores para las dos condiciones (de hecho, el conductor no sabe qué tipo de gasolina se utiliza en las pruebas). Con los datos de rendimiento de la tabla adjunta, probar la hipótesis de que no existe diferencia entre el kilometraje promedio que se obtiene con o sin aditivo, utilizando un nivel de significación del 5%.

## RENDIMIENTO DE KILOMETRAJE

| Automóvil | Kilometraje con aditivo | Kilometraje sin aditivo |
|---|---|---|
| 1 | 12,66 | 12,48 |
| 2 | 12,34 | 12,31 |
| 3 | 11,00 | 11,14 |
| 4 | 10,10 | 10,21 |
| 5 | 9,79 | 9,69 |
| 6 | 8,86 | 8,90 |
| 7 | 8,84 | 8,24 |
| 8 | 7,79 | 7,59 |
| 9 | 7,55 | 7,41 |
| 10 | 7.00 | 6,90 |

**EJEMPLO 11:**

La vida útil promedio de una muestra aleatoria de n = 10 focos es de $\overline{x}$ = 4.000 horas, con S = 200 horas. Suponiendo que la vida útil de los focos tiene una distribución normal, y que antes de obtener la muestra se plantea la hipótesis de que $\sigma$ no es superior a 150 horas. Con base a los resultados muestrales probar dicha hipótesis a un nivel de significación del 1%.

**EJEMPLO 12:**

Con los datos del **ejemplo 11**, en otra muestra B de focos, para los cuales se supone que tienen una vida útil con distribución normal, se obtiene una muestra aleatoria $n_B$ = 8, obteniéndose que $\overline{x}_B$ = 4.300 horas y $S_B$ = 250 horas. Probar la hipótesis de que las muestras se obtuvieron con varianzas iguales al nivel de significación del 10%.

## PRUEBAS DE HIPÓTESIS: PROBLEMAS

### PROBLEMA 1:

El gerente de crédito de una cadena de tiendas afirma que el saldo mensual promedio de los tarjeta-habitantes es de 30$. Para probar su afirmación, un auditor selecciona una muestra aleatoria de 100 cuentas y encuentra que el saldo promedio es de 35$, con una desviación estándar de la muestra de 12,50$. Con un nivel de significación de 0,01, ¿a qué conclusiones llegaría el auditor?

### PROBLEMA 2:

El fabricante de los acumuladores Dillco asevera que los acumuladores Dillco duran un promedio de 40.000 millas. Se probó una muestra aleatoria de 81 acumuladores y se tuvo como resultado una duración promedio de 40.800 millas, con una desviación estándar de la muestra de 5.400 millas. Con un nivel de significación de 0,05, ¿es verdadera la afirmación del fabricante de los acumuladores Dillco?

### PROBLEMA 3:

Aceros Pimiango S.A. fabrica barras de acero. El proceso de producción hace barras con una longitud promedio de al menos 2,8 pies cuando el proceso funciona correctamente. Se selecciona una muestra de 25 barras en la línea de producción. La muestra indica una longitud promedio de 2,43 pies y una desviación estándar de 0,20 pies. La compañía desea determinar si la máquina necesita algún ajuste.

a) Indique las hipótesis nula y alternativa.

b) Si la compañía desea probar la hipótesis al nivel de significación de 0,05,¿qué decisión tomaría?

**PROBLEMA 4:**

El propietario de una distribuidora de línea blanca al mayoreo debe estimar su promedio de cuentas por cobrar al final de un período de un mes. Con base en la experiencia pasada, estima que el promedio de cuentas por cobrar es de 10.000$. En vez de aceptar esta estimación del propietario, el contable decide seleccionar una muestra aleatoria de 36 cuentas. Los resultados son los siguientes $\overline{X}$ = 12.300$ y S = 3.300$. Con un nivel de significación de 0,10, ¿se debería aceptar la estimación del propietario?

**PROBLEMA 5:**

El gerente de personal de una empresa querría determinar la cantidad de tiempo que necesitan sus empleados para llegar a su trabajo. Se selecciona una muestra aleatoria de 12 empleados y se registra el tiempo en minutos para llegar al trabajo, con los siguientes resultados:

15 30 50 60 25 65 45 90 75 50 50 20

a) Con un nivel de significación de 0,01, ¿hay pruebas de que el tiempo promedio de viaje de los empleados es de menos de 60 minutos?

b) ¿Qué suposición es necesaria para efectuar esta prueba?

**PROBLEMA 6:**

Una agencia de informes de consumidores deseaba determinar si una calculadora de ``marca desconocida'' se vende a un precio más bajo que una calculadora de ``marca famosa'' del mismo tipo. Se seleccionó una muestra aleatoria de 8 tiendas y se registraron los precios en las tiendas de cada una de las calculadoras, con los siguientes resultados:

| Tienda | Marca Desconocida | Marca Famosa |
|---|---|---|
| 1 | 10 | 11 |
| 2 | 8 | 11 |
| 3 | 7 | 10 |
| 4 | 9 | 12 |
| 5 | 11 | 11 |
| 6 | 10 | 13 |
| 7 | 9 | 12 |
| 8 | 8 | 10 |

a) Al nivel de significación de 0,01, ¿se vende la marca desconocida a un precio más bajo?

b) ¿Qué suposición es necesaria para efectuar esta prueba?

**PROBLEMA 7:**

Una máquina vendedora de refrescos está proyectada para despachar, cuando funciona correctamente, cuando menos 7 onzas de refresco por vaso, con una desviación estándar de 0,2 onzas. Si el estadístico selecciona una muestra aleatoria

de 16 vasos para un estudio especial y el estadístico está dispuesto a tener un error de tipo I ($\alpha$) del 5%, calcule la potencia de la prueba y la probabilidad de un error de tipo II ($\beta$) si la cantidad promedio de población despachada es:

a) 6,9 onzas por vaso.

b) 6,8 onzas por vaso.

c) Si el estadístico desea tener una potencia del 99% para detectar un desplazamiento en la media de la población de 7,0 onzas a 6,9 onzas, ¿qué tamaño de muestra debe seleccionar?

**PROBLEMA 8:**

Con los datos del **ejercicio 7**, si el estadístico desea tener un riesgo $\alpha$ del 1%, calcule la potencia de la prueba y la probabilidad de un error de tipo II ($\beta$) si la cantidad promedio de población despachada es:

a) 6,9 onzas.

b) 6,8 onzas.

c) Compare los resultados en **a)** y **b)** de este ejercicio y del **ejercicio 7**.

**PROBLEMA 9:**

El director de presupuestos de una empresa deseaba determinar si había alguna diferencia en las cuentas de gastos de representación de los ejecutivos de dos departamentos de la empresa. Se seleccionó una muestra

aleatoria de 15 cuentas de gastos del Departamento 1 y 15 cuentas del Departamento 2, con los siguientes resultados:

| Departamento 1 | Departamento 2 |
| --- | --- |
| $x_1$ = 27,50 $ | $x_2$ = 23,25 $ |
| $S_1$ = 6,00 $ | $S_2$ = 7,50 $ |

Al nivel de significación de 0,01, ¿es más alto el gasto de representación promedio en el Departamento 1?

**PROBLEMA 10:**

El gerente de marketing de una cadena de tiendas de flores y plantas cree que habrá una gran demanda de bulbos de gladiolo en las tiendas de los suburbios durante el mes de mayo. Si las ventas de bulbos exceden de 100$ por semana, se venderán en todas las tiendas suburbanas de la cadena. Se selecciona una muestra aleatoria de 16 tiendas y los resultados de la prueba arrojaron un promedio de ventas de 120$, con una desviación estándar de la muestra de 25$.

a) Para un nivel de significación del 0,01, ¿se deberían vender los bulbos de gladiolo en todas las tiendas suburbanas de la cadena?

b) ¿Qué suposición es necesaria para efectuar esta prueba?

**PROBLEMA 11:**

Un analista de sistemas está probando la posibilidad de usar un nuevo sistema de computadora. El analista cambiará el procesamiento al nuevo sistema sólo si hay pruebas de que el nuevo sistema usa menos tiempo para el procesamiento que el sistema antiguo. A fin de tomar una decisión, se seleccionó una muestra de siete trabajos y se registró el tiempo de procesamiento, en segundos, en los dos sistemas, con los siguientes resultados:

| Trabajos | Viejo | Nuevo |
|---|---|---|
| 1 | 8 | 6 |
| 2 | 4 | 3 |
| 3 | 10 | 7 |
| 4 | 9 | 8 |
| 5 | 8 | 5 |
| 6 | 7 | 8 |
| 7 | 12 | 9 |

a) Al nivel de significación del 0,01, ¿usa menos tiempo el nuevo sistema para el procesamiento?

b) ¿Qué suposición es necesaria para efectuar esta prueba?

SOLUCIONES

1) $H_o$: $\mu_x = 30$. Rechazamos $H_o$.

2) $H_o$: $\mu_x = 40.000$. Aceptamos $H_o$.

3) $H_o$: $\mu_x \geq 2,8$. Rechazamos $H_o$.

4) $H_o$: $\mu_x = 10.000$. Rechazamos $H_o$.

5) $H_o$: $\mu_x \geq 60$. Aceptamos $H_o$.

6) $H_o$: $\mu_d \leq 0$. Rechazamos $H_o$.

7) $H_o$: $\mu_x \geq 7$; $\beta_{6,9} = 0,3612$; $(1 - \beta) = 0,6388$;

$(1 - \beta_{6,8}) = 0,9907$.

8) $H_o$: $\mu_x \geq 7$; $(1 - \beta_{6,9}) = 0,3707$;

$(1 - \beta_{6,8}) = 0,9529$. Si $\alpha \downarrow \Rightarrow \beta \uparrow \Rightarrow (1 - \beta) \downarrow$.

9) $H_o$: $\sigma_x^2 = \sigma_y^2$. Aceptamos $H_o$.

10) $H_o$: $\mu_x \leq 100$. Rechazamos $H_o$.

11) $H_o$: $\mu_d = 0$. Aceptamos $H_o$.

## PRUEBAS DE HIPÓTESIS SOBRE LA $\mu$ DE LA POBLACIÓN

## PRUEBAS BILATERALES

**POBLACIÓN**

Nat: Normal o n > 30 TCL

$\mu$: ?

$\sigma_x^2$: CONOCIDA

* $H_o$: $\mu = \mu_o$

  $H_1$: $\mu \neq \mu_o$

* El EC es: $\dfrac{\overline{X} - \mu_0}{\sigma_{X/\sqrt{n}}} = Z$

* **Valores críticos:** $+Z_{\alpha/2}$; $-Z_{\alpha/2}$

* **Decisión:** Si EC $\geq +Z_{\alpha/2}$
  EC $\leq -Z_{\alpha/2}$ > **RECHAZO $H_o$**

# PRUEBAS DE HIPÓTESIS SOBRE LA $\mu$ DE LA POBLACIÓN

## PRUEBAS UNILATERALES

**POBLACIÓN**

Nat: Normal o $n > 30$ TCL

$\mu_x$: ?

$\sigma_x^2$: CONOCIDA

* $H_o: \mu \leq \mu_o$ o $H_o: \mu = \mu_o$
  $H_1: \mu > \mu_o$ $H_1: \mu > \mu_o$

* $H_o: \mu \geq \mu_o$ o $H_o: \mu = \mu_o$
  $H_1: \mu < \mu_o$ $H_1: \mu < \mu_o$

* **EC:** $$\frac{\overline{X} - \mu_0}{\sigma_{X/\sqrt{n}}} = Z$$

* **Valor crítico:** $+Z_\alpha$ o $-Z_\alpha$

* **Decisión:** Si $EC > +Z_\alpha$ **RECHAZO $H_o$**

  **o**

  Si $EC < -Z_\alpha$ **RECHAZO $H_o$**

# PRUEBAS DE HIPÓTESIS SOBRE LA $\mu$ DE LA POBLACIÓN

## PRUEBAS BILATERALES

**POBLACIÓN**

Nat: Normal o n > 30 TCL

$\mu$: ?

$\sigma_x^2$: DESCONOCIDA

* $\hat{\sigma}_x^2 = S^2$

* $H_o: \mu = \mu_o$

  $H_1: \mu \neq \mu_o$

* El EC es: $\dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} = t_{n-1}$

* **Valores críticos:** $t_{n-1;\ \alpha/2}$; $-t_{n-1;\ \alpha/2}$

* **Decisión:** Si $EC \geq t_{n-1;\ \alpha/2}$

  $EC \leq -t_{n-1;\ \alpha/2}$

  > **RECHAZO $H_o$**

## PRUEBAS DE HIPÓTESIS SOBRE LA $\mu$ DE LA POBLACIÓN

## PRUEBAS UNILATERALES

**POBLACIÓN**

Nat: Normal o n > 30 TCL

$\mu$: ?

$\sigma_x^2$: DESCONOCIDA

* $\hat{\sigma}_x^2 = S^2$

* $H_o: \mu \geq \mu_o$ o $H_o: \mu = \mu_o$

  $H_1: \mu < \mu_o$ $H_1: \mu < \mu_o$

* $H_o: \mu \leq \mu_o$ o $H_o: \mu = \mu_o$

  $H_1: \mu > \mu_o$ $H_1: \mu > \mu_o$

* **EC:** $\dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} = t_{n-1}$

* **Valor crítico:** $-t_{n-1;\,\alpha}$ o $t_{n-1;\,\alpha}$

* **Decisión:** Si $EC \leq -t_{n-1;\,\alpha}$ **RECHAZO $H_o$**

  o

  Si $EC \geq t_{n-1;\,\alpha}$ **RECHAZO $H_o$**

# PRUEBAS DE HIPÓTESIS SOBRE LAS $\mu$ DE DOS POBLACIONES

## PRUEBAS BILATERALES

**POBLACIONES**

* Nat X: Normal o $n_x > 30$
* Nat Y: Normal o $n_y > 30$

> INDEPENDIENTES

* $\sigma_x^2$ ; $\sigma_y^2$ CONOCIDAS

* $H_0: \mu_x = \mu_y$ o $H_0: \mu_x - \mu_y = 0$

  $H_1: \mu_x \neq \mu_y$ $H_1: \mu_x - \mu_y \neq 0$

* **EC:** $$\frac{(\overline{X} - \overline{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} = Z$$

* **Valores críticos:** $+Z_{\alpha/2}$; $-Z_{\alpha/2}$

* **Decisión:** Si $EC \geq +Z_{\alpha/2}$

  $EC \leq -Z_{\alpha/2}$

> **RECHAZO $H_o$**

## PRUEBAS DE HIPÓTESIS SOBRE LAS $\mu$ DE DOS POBLACIONES

## PRUEBAS UNILATERALES

* $H_o: \mu_x \le \mu_y$ o $H_o: \mu_x \ge \mu_y$

  $H_1: \mu_x > \mu_y$  $H_1: \mu_x < \mu_y$

* EC: $$\frac{(\overline{X} - \overline{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} = Z$$

* Valores críticos: $+Z_\alpha$ o $-Z_\alpha$

* Decisión: Si EC $\ge Z_\alpha$ RECHAZO $H_o$

  o

  Si EC $\le -Z_\alpha$ RECHAZO $H_o$

## PRUEBAS DE HIPÓTESIS SOBRE LAS μ DE DOS POBLACIONES

## PRUEBAS BILATERALES

**POBLACIONES**

* Nat X: Normal o $n_x > 30$
* Nat Y: Normal o $n_y > 30$

} INDEPENDIENTES

* $\sigma_x^2 = \sigma_y^2$ DESCONOCIDAS

* $H_o: \mu_x = \mu_y$ o $H_o: \mu_x - \mu_y = 0$

  $H_1: \mu_x \neq \mu_y$ $H_1: \mu_x - \mu_y \neq 0$

* **EC:** $$\frac{(\overline{X} - \overline{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{S_p^2}{n_x} + \frac{S_p^2}{n_y}}} = t_{n_x + n_y - 2}$$

$$S_p^2 = \frac{(n_x - 1)\, S_x^2 + (n_y - 1)\, S_y^2}{n_x + n_y - 2}$$

* **Valores críticos:** $-t_{\alpha/2}$; $+t_{\alpha/2}$

* **Decisión:** Si $EC \geq +t_{\alpha/2}$

  $EC \leq -t_{\alpha/2}$

} **RECHAZO $H_o$**

## PRUEBAS DE HIPÓTESIS SOBRE LAS μ DE DOS POBLACIONES

## PRUEBAS UNILATERALES

* $H_o: \mu_x \leq \mu_y$ o $H_o: \mu_x \geq \mu_y$

  $H_1: \mu_x > \mu_y$ $H_1: \mu_x < \mu_y$

* EC: $$\frac{(\overline{X} - \overline{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{S_p^2}{n_x} + \frac{S_p^2}{n_y}}} = t_{n_x + n_y - 2}$$

* **Valores críticos:** $+t_{\alpha;\ n_x + n_y - 2}$ o $-t_{\alpha;\ n_x + n_y - 2}$

* **Decisión:** Si EC $\geq t_{\alpha;\ n_x + n_y - 2}$ **RECHAZO $H_o$**

  **o**

  Si EC $\leq -t_{\alpha;\ n_x + n_y - 2}$ **RECHAZO $H_o$**

## PRUEBAS DE HIPÓTESIS SOBRE LAS μ DE DOS POBLACIONES

## PRUEBAS BILATERALES

**POBLACIONES**

* Nat X: Normal o $n_x > 30$
* Nat Y: Normal o $n_y > 30$

> INDEPENDIENTES

* $\sigma_x^2 \neq \sigma_y^2$ DESCONOCIDAS

* $H_o: \mu_x = \mu_y$ o $H_o: \mu_x - \mu_y = 0$

  $H_1: \mu_x \neq \mu_y$ o $H_1: \mu_x - \mu_y \neq 0$

* EC:

$$\frac{(\bar{X} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}} = t_K$$

$$K = \frac{(n_x - 1)(n_y - 1)}{(n_y - 1)C^2 + (1 - C)^2 (n_x - 1)}$$

$$C = \frac{\frac{S_x^2}{n_x}}{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}$$

* Valores críticos: $-t_{K,\ \alpha/2}$; $+t_{K,\ \alpha/2}$

* Decisión: Si $EC \geq +t_{K,\ \alpha/2}$

  $EC \leq -t_{K,\ \alpha/2}$

> RECHAZO $H_o$

## PRUEBAS DE HIPÓTESIS SOBRE LAS $\mu$ DE DOS POBLACIONES

## PRUEBAS UNILATERALES

* $H_o: \mu_x \leq \mu_y$ o $H_o: \mu_x \geq \mu_y$

  $H_1: \mu_x > \mu_y$ $H_1: \mu_x < \mu_y$

* EC: $$\frac{(\overline{X} - \overline{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}} = t_K$$

* Valores críticos: $+t_{\alpha;\,K}$ o $-t_{\alpha;\,K}$

* Decisión: Si EC $\geq +t_{\alpha;\,K}$ RECHAZO $H_o$

  o

  Si EC $\leq -t_{\alpha;\,K}$ RECHAZO $H_o$

# PRUEBAS DE HIPÓTESIS SOBRE LAS $\mu$ DE DOS POBLACIONES DEPENDIENTES

## PRUEBAS BILATERALES

**POBLACIONES**

* Nat X: Normal o $n_x > 30$
* Nat Y: Normal o $n_y > 30$

> DEPENDIENTES

* $\sigma_x^2$ ; $\sigma_y^2$ CONOCIDAS o DESCONOCIDAS

* Hacemos d = X - Y

* $H_o$: $\mu_x = \mu_y$ o $\mu_d = 0$ ($\mu_x - \mu_y = 0$)

  $H_1$: $\mu_x \neq \mu_y$ o $\mu_d \neq 0$

* **EC:** $$\frac{\bar{d} - \mu_d}{\sqrt{\frac{S^2 d}{n}}} = t_{n-1}$$ $n_x = n_y = n$

* **Valores críticos:** $-t_{n-1;\ \alpha/2}$; $+t_{n-1;\ \alpha/2}$

* **Decisión:** Si $EC \geq t_{n-1;\ \alpha/2}$

  $EC \leq -t_{n-1;\ \alpha/2}$

  > **RECHAZO $H_o$**

## PRUEBAS DE HIPÓTESIS SOBRE LAS μ DE DOS POBLACIONES DEPENDIENTES

## PRUEBAS UNILATERALES

* $H_o$: $\mu_d \leq 0$ o $H_o$: $\mu_d \geq 0$

  $H_1$: $\mu_d > 0$ $\quad$ $H_1$: $\mu_d < 0$

* **EC:** $\dfrac{\bar{d} - \mu_d}{\sqrt{\dfrac{S^2 d}{n}}} = t_{n-1}$

* **Valores críticos:** $t_{n-1;\,\alpha}$ ó $-t_{n-1;\,\alpha}$

* **Decisión:** Si EC $\geq t_{n-1;\,\alpha}$ **RECHAZO $H_o$**

  o

  Si EC $\leq -t_{n-1;\,\alpha}$ **RECHAZO $H_o$**

## PRUEBAS DE HIPÓTESIS SOBRE LA VARIANZA POBLACIONAL

## PRUEBAS BILATERALES

* $H_o$: $\sigma^2 = \sigma_0^2$

  $H_1$: $\sigma^2 \neq \sigma_0^2$

* **EC:** $\frac{(n - 1) S^2}{\sigma_0^2} = X_{n-1}^2$

* **Valores críticos:** $X_{1-\alpha/2}^2$; $X_{\alpha/2}^2$ (con [n - 1] g.l.)

* **Decisión:** Si EC $\geq X_{\alpha/2}^2$
  EC $\leq X_{1-\alpha/2}^2$
  > **RECHAZO $H_o$**

# PRUEBAS DE HIPÓTESIS SOBRE LA VARIANZA POBLACIONAL

## PRUEBAS UNILATERALES

**1º Caso**

* $H_o: \sigma^2 \leq \sigma_0^2$ o $H_o: \sigma^2 = \sigma_0^2$
  $H_1: \sigma^2 > \sigma_0^2$ $H_1: \sigma^2 > \sigma_0^2$

* **EC:** $\frac{(n - 1) S^2}{\sigma_0^2} = X_{n-1}^2$

* **Valor crítico:** $X^2_{(n-1), \alpha}$

* **Decisión:** Si EC $\geq X^2_{(n-1), \alpha}$ **RECHAZO $H_o$**

**2º Caso**

* $H_o: \sigma^2 \geq \sigma_0^2$ o $H_o: \sigma^2 = \sigma_0^2$
  $H_1: \sigma^2 < \sigma_0^2$ $H_1: \sigma^2 < \sigma_0^2$

* **EC:** $\frac{(n - 1) S^2}{\sigma_0^2} = X_{(n-1)}^2$

* **Valor crítico:** $X^2_{(n-1), (1-\alpha)}$

* **Decisión:** Si EC $\leq X^2_{(n-1), (1-\alpha)}$ **RECHAZO $H_o$**

## PRUEBAS DE HIPÓTESIS SOBRE LA VARIANZA DE DOS POBLACIONES

* $H_o$: $\sigma_x^2 = \sigma_y^2$

  $H_1$: $\sigma_x^2 \neq \sigma_y^2$

* **EC:** $\dfrac{S_x^2}{S_y^2} = F_{n_x - 1;\, n_y - 1}$

* **Valores críticos:**

  $F_{1 - \alpha/2;\, n_x - 1;\, n_y - 1}$

  $F_{\alpha/2;\, n_x - 1;\, n_y - 1}$

* **Decisión:**

  Si EC $\geq F_{\alpha/2;\, n_x - 1;\, n_y - 1}$
  
  **o**
  
  EC $\leq F_{1 - \alpha/2;\, n_x - 1;\, n_y - 1}$
  
  > **RECHAZO $H_o$**

## PRUEBAS DE HIPÓTESIS PARA LA PROPORCIÓN DE ÉXITOS

## PRUEBAS BILATERALES

**POBLACIÓN**

Nat: Binomial

* $H_o: P = P_o$

  $H_1: P \neq P_o$

* n suficientemente grande: Aprox. por Dist. Normal

* **EC:** $$\frac{(\bar{P} - P) \pm 0'5}{\sqrt{\frac{P\,(1 - P)}{n}}} = Z$$

* **Valores críticos:** $+Z_{\alpha/2}$; $-Z_{\alpha/2}$

* **Decisión:** Si $EC \geq +Z_{\alpha/2}$ | $EC \leq -Z_{\alpha/2}$ } **RECHAZO $H_o$**

## PRUEBAS DE HIPÓTESIS PARA LA PROPORCIÓN DE ÉXITOS

## PRUEBAS UNILATERALES

* $H_o$: $P \leq P_o$ o $H_o$: $P \geq P_o$

  $H_1$: $P > P_o$ $H_1$: $P < P_o$

* Si **n** suficientemente grande: Aprox. mediante Dist. Normal

* **EC:** $$\frac{(\overline{P} - P) \pm 0'5}{\sqrt{\frac{P\,(1 - P)}{n}}} = Z$$

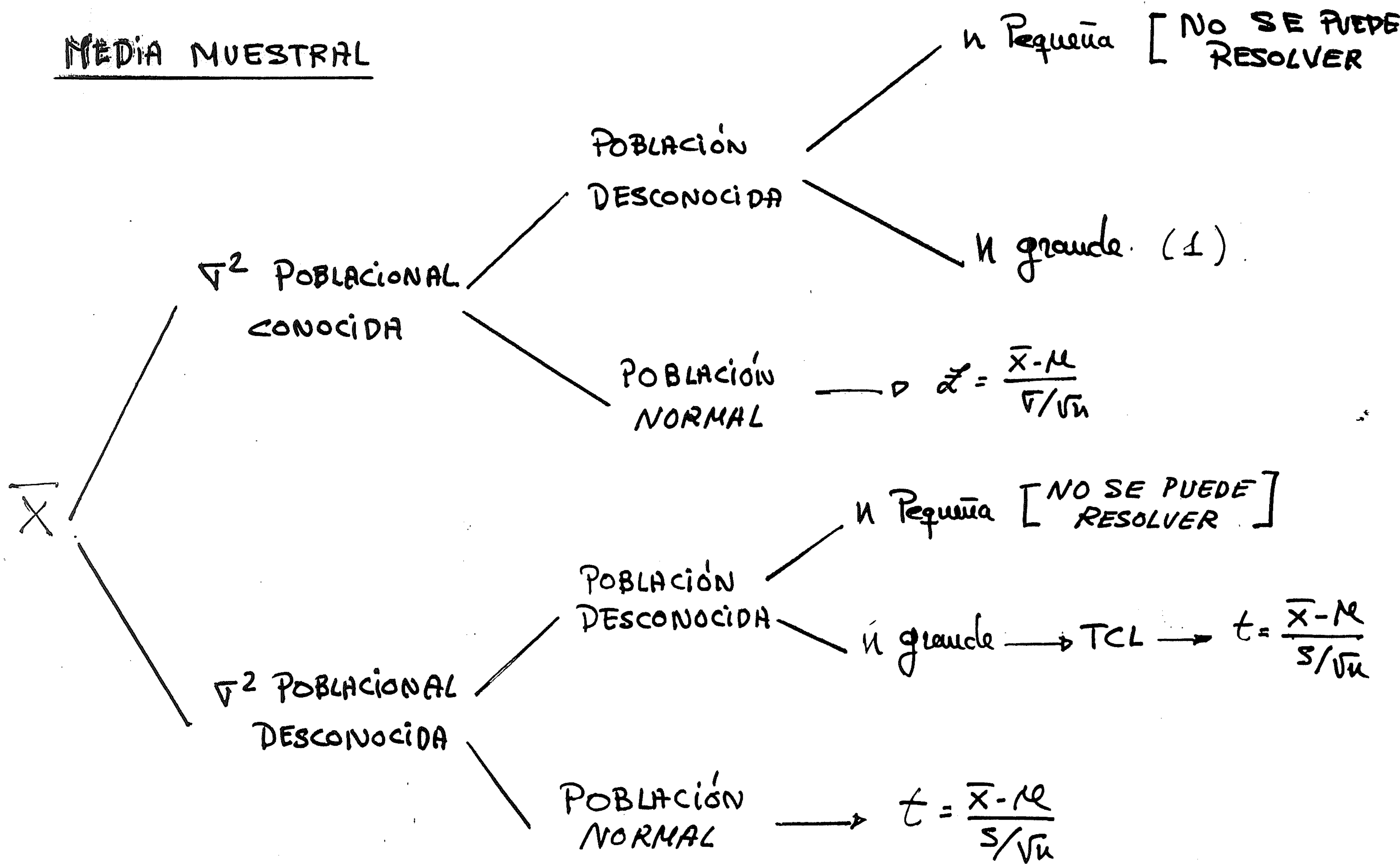
* **Valores críticos:** $+Z_\alpha$ o $-Z_\alpha$

* **Decisión:** Si $EC > +Z_\alpha$ **RECHAZO $H_o$**

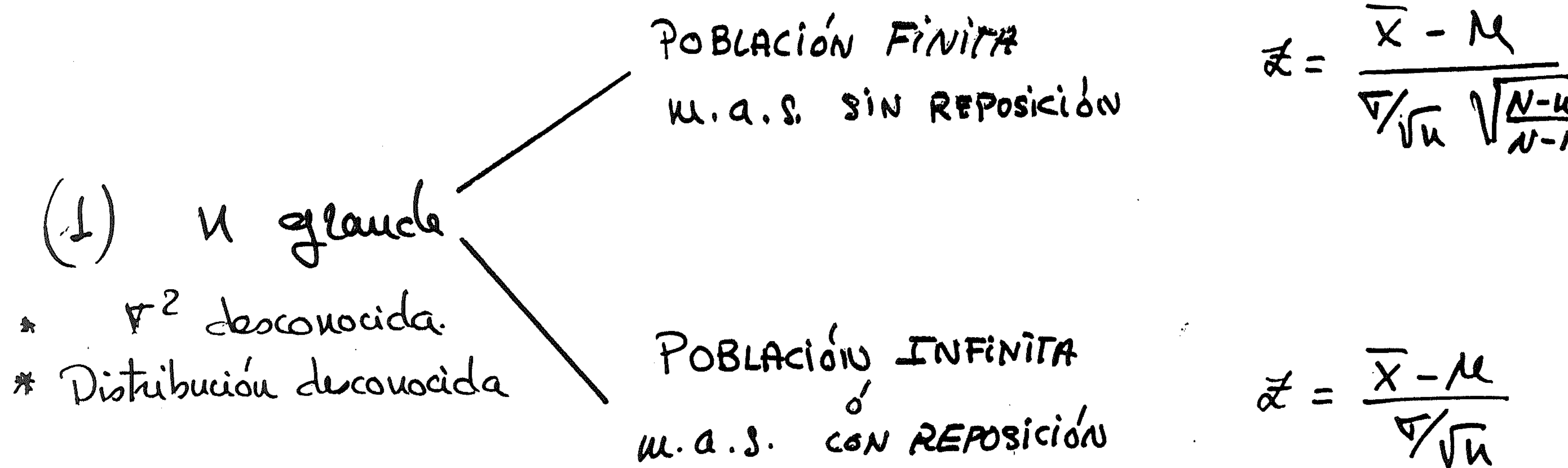
  **o**

  Si $EC < -Z_\alpha$ **RECHAZO $H_o$**

# MEDIA MUESTRAL



# MEDIA MUESTRAL



(1) $n$ grande

* $\sigma^2$ desconocida.
* Distribución desconocida

POBLACIÓN FINITA
m.a.s. SIN REPOSICIÓN

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}\,\sqrt{\frac{N-n}{N-1}}}$$

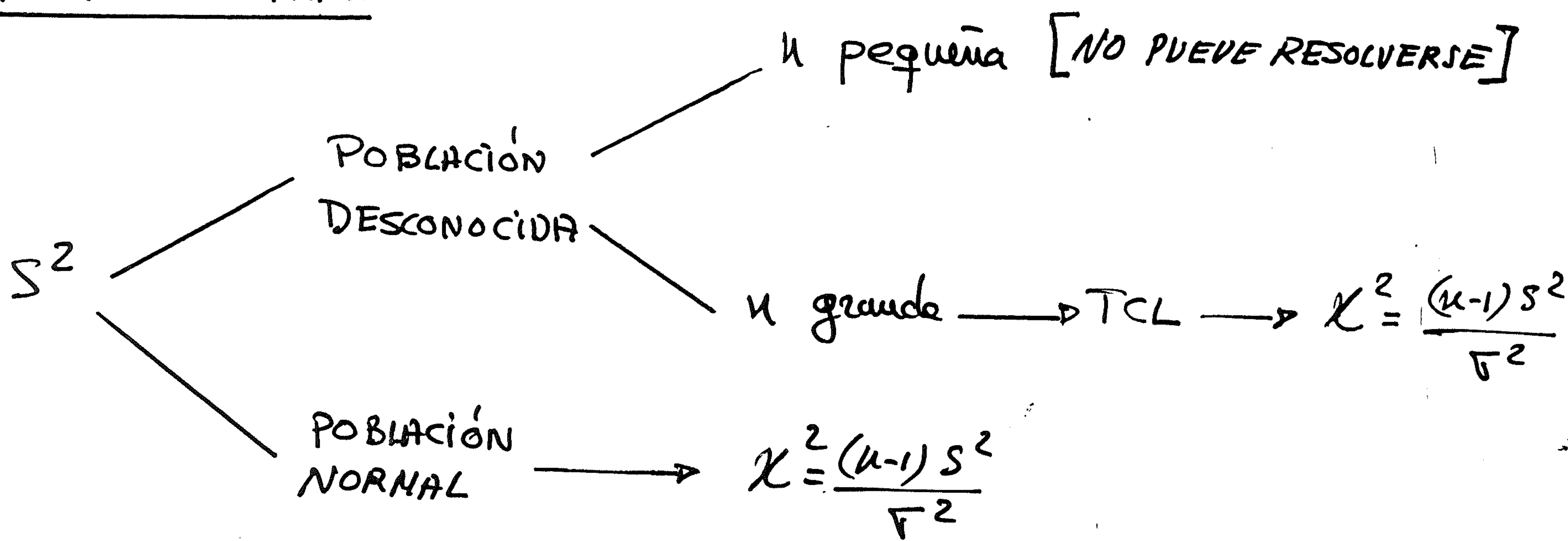
POBLACIÓN INFINITA
ó
m.a.s. CON REPOSICIÓN

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

¿CUANDO SE CONSIDERA $n$ grande?

* SI LA DISTRIBUCIÓN ES SIMETRICA $n \geq 10$
* SI NO LO ES $n \geq 30$

## VARIANZA MUESTRAL




## RAZÓN DE VARIANZAS MUESTRALES

$$\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} \longrightarrow F_{n_X-1;\, n_Y-1}$$

POBLACIÓN NORMAL.

X e Y INDEPENDIENTES

# DIFERENCIA DE MEDIAS MUESTRALES

$\bar{X} - \bar{Y}$

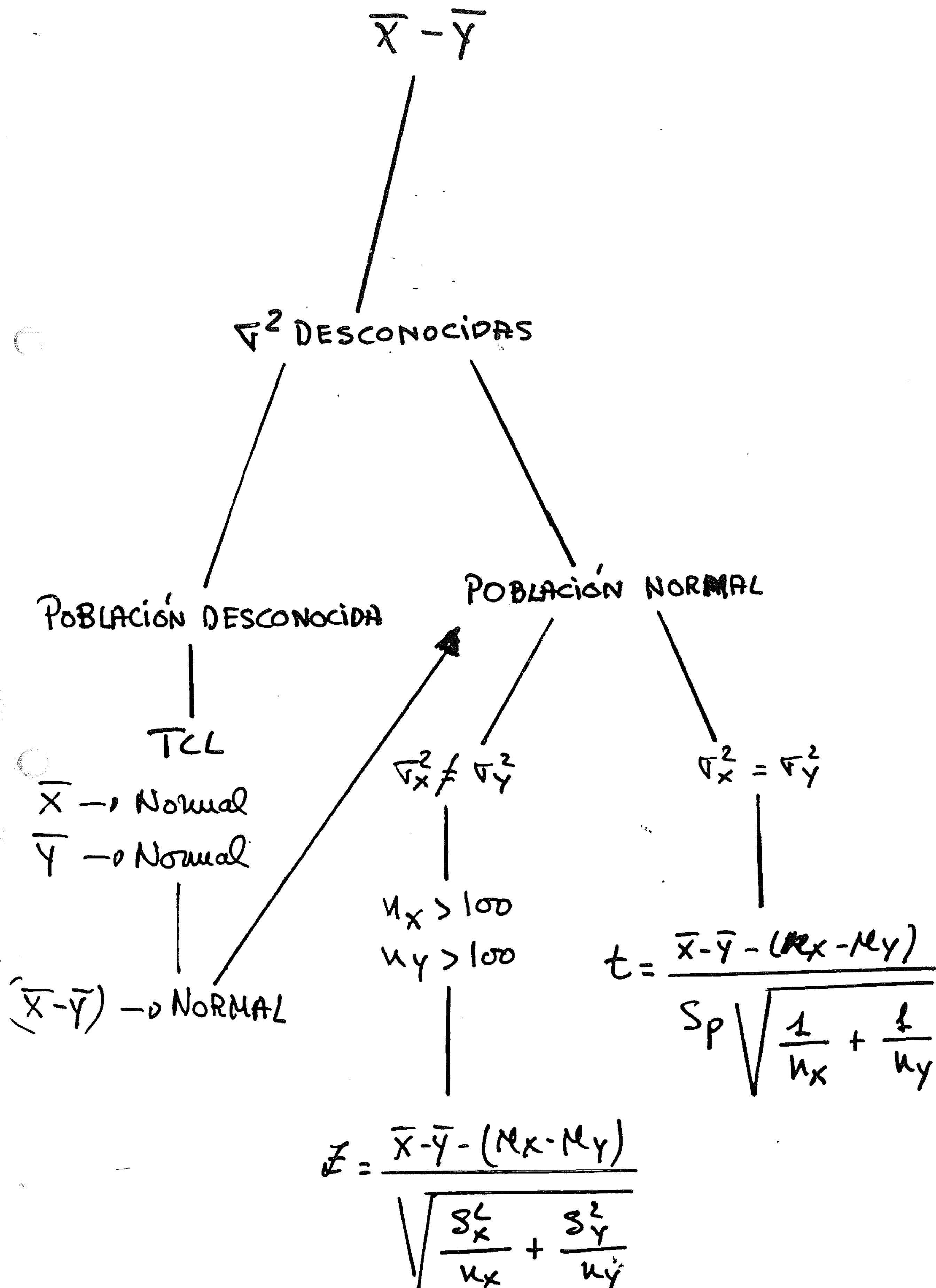
- $\sigma^2$ DESCONOCIDAS (2)
- $\sigma^2$ CONOCIDAS
  - POBLACIÓN DESCONOCIDA
    - TCL
    - $\bar{X}$ → NORMAL
    - $\bar{Y}$ → NORMAL
    - $(\bar{X}-\bar{Y})$ → NORMAL (→ POBLACIÓN NORMAL)
  - POBLACIÓN NORMAL
    - $\sigma_X^2 \neq \sigma_Y^2$

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}}$$

    - $\sigma_X^2 = \sigma_Y^2$

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sigma \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}}$$

(2) DISTRIBUCIÓN DE LA DIFERENCIA DE MEDIAS MUESTRALES



Iratxe Berriozabal Campillo.

24.10.96.

## EJEMPLO 15

Se sabe que en San Sebastian el 30 % de la población son consumidores de café. Si mediante m.a.s. tomamos una muestra de 4 ciudadanos de San Sebastian, se pide :

a) La distribución muestral del nº de consumidores de cafe por muestra.

b) ¿ Cual es la probabilidad de que en una muestra tengamos 3 o más consumidores de cafe ?.

c) Utilizando la aproximación de la distribució Normal, calcula el apartado b).

d) Comparar los resultados de los apartados b) y c)

Iratxe Semiozabal Campillo.
05.11.96.

PROBLEMA :

Se quiere estudiar el nº de niños nacidos con una determinada caracteristica en dos poblaciones A y B . En la población A se han tenido encuenta los ultimos 300 nacimientos, encontrandose que en 45 de ellos se daban estas caracteristicas, mientras que en la población B se han tenido encuenta los 350 ultimos nacimientos, encontrandose 57 casos en los que se daban estas caracteristicas.

Definimos un indice C como :

$$C = \left(\frac{X_A - \mu_A}{\sigma_A}\right)^2 + \left(\frac{X_B - \mu_B}{\sigma_B}\right)^2$$

Teniendo encuenta que la probabilidad P de que un niño tenga la caracteristica correspondiente es 0,15 . ¿ Cual sera la probabilidad de que C sea mayor que el obtenido con los valores $X_A$ y $X_B$ observados ?.

PROBLEMA:

$P[c > c_1]$

$$c_1 = \left(\frac{x_A - \mu_A}{\sigma_A}\right)^2 + \left(\frac{x_B - \mu_B}{\sigma_B}\right)^2$$

$x_A$: 45 $\mu_A$: la media del nº de nacimientos con esa característica A en esos
$x_B$: 57 300 nacimientos

Por hablar de una $\mu$ y $\sigma$ estaremos hablando de una v.a.

v.a: niños nacidos con esa característica. Sigue una distribución binomial de parámetros. $n = 300$
$p = 0'15$

v.a. $A \rightarrow B(n = 300; p = 0'15)$.

Es binomial porque: - tenemos 'n' pruebas independientes.
- sólo hay dos resultados posibles
- el valor de "p" es cte.

$\mu_A = n.p = 300.0'15 = 45$
$\sigma_A = \sqrt{n.p.q} = \sqrt{300.0'15.0'85} = 6'18$ $0'85 = 1 - 0'15$

v.a. $B \rightarrow B(n = 350; p = 0'15)$
$\mu_B = n.p = 350.0'15 = 52'5$
$\sigma_B = \sqrt{n.p.q} = \sqrt{350.0'15.0'85} = 6'68$

Sustituimos esos valores en la fórmula de arriba y obtenemos que:
$P[c > 0'4538] =$
$c = \frac{x_A - \mu_A}{\sigma_A}$; al tipificar, obtengo una puntuación $z \rightarrow N(0,1)$, si la elevo al cuadrado y le sumo otra puntuación $z^2$ obtengo una puntuación $\chi^2$